图书情报工作
第64卷 第23期 2020年12月



# 多媒体网络舆情信息并发获取模型构建及实证分析*

■ 许烨婧　黄微　赵江元
吉林大学管理学院　长春 130022

摘　要：[目的/意义]网络舆情信息并发获取模型构建及实证研究有助于从海量的舆情信息中及时获取所需关键性信息，为其有效分析提供数据保障。[方法/过程]通过对当前舆情信息获取研究现状的综合分析，明确多媒体网络舆情信息并发获取的模型构成要素，融合 DEMATEL、AHP、FMF 三种数理分析法进行模型构建并据此展开实证分析。[结果/结论]研究结果表明所得数据结论与舆情事件客观情况较为相符，可以作为舆情信息并发获取判断的依据。

关键词：多媒体　网络舆情　并发获取　模型构建
分类号：G203
DOI：10.13266/j.issn.0252－3116.2020.23.012



## 1　引言

互联网科学技术的急速变革与无边界发展，使得信息呈现海量倍速增长趋势。截至2020年3月，中国互联网络信息中心（CNNIC）发布的《第45次中国互联网络发展状况统计报告》显示[1]，我国网民规模达到9.04亿，互联网普及率为64.5%，较2018年底增幅为4.9%。种种迹象表明，信息源数量之大，传播途径之多，扩散速度之快前所未有，尤其是伴随网络环境的复杂性、随机性、多变性，网络舆情的广博覆盖映射了社情民意，深刻影响着民众的态度、意识。如何透过纷繁复杂的舆情信息有效获取识别关键信息成为重中之重，特别是网络舆情信息所具有的巨量性、随机性、内隐性、突变性等特点，极大增加了舆情信息有效获取的难度，虽然目前网络舆情信息采集、获取的相关研究具有一定规模，但当前有关网络舆情并发获取的相关研究极少，因此本文主要从并发与并行处理、信息获取的视角展开研究现状分析。在并发处理的效率方面，J. Xue等[2]介绍了高并行处理系统，当存在多并发操作时，能够实现数据共享，扩展逻辑调度空间，增强数据并发处理性能，减少容错率。在并行设计方面，H. Zheng等[3]通过展开并行设计活动，发现存在的问题，经过深层次探究，明确其影响因素，在此基础上构建结构关系模型。信息获取模型构建方面，S. Sun等提出了一种分布式增量信息获取模型，有助于及时从大规模文本数据中发现和获取互联网上的增量数据[4]。信息获取方法上，Y. Liu等提出了基于概念格和包含度的有序信息系统知识获取方法[5]。J. Liu[6]以公共部门网站为研究对象，对其信息采集获取原则、渠道及常用方法展开了分析与探讨。在并发处理模型构建方面，李辉等[7]以大数据平台的数据为对象，构建并发处理模型，并通过仿真实验的方式进行验证。洪亮等[8]构建系统动力模型并采用模拟仿真实验法验证，结果表明网络舆情与网民、事件、媒体等互动性强，其指出了相关影响因素并提供了应对策略。朱红灿等[9]以心流理论为基础，分析信息获取的影响因素，建构使用意图模型，通过实证分析法验证模型的合理性。在信息获取测量方面，靳江艳等[10]构建了信息多色集合的层次模型，确立了信息获取关系矩阵，以实证研究的方式验证模型的可操作性。信息采集方面，李莉等聚焦网络爬虫，将其划分为网页爬取、预处理、关联性评价3个模块，通过实验证明该方法能够取得良好的查准率[11]。张洪豪等模拟浏览器向所需Web页面发送虚拟请求，实现了一种基于Web爬虫的网页信息获取系统[12]。

多媒体网络舆情具有明显的内隐与外现双性特

* 本文系国家社会科学基金重大项目“大数据驱动的社交网络舆情主题图谱构建及调控策略研究”（项目编号：18ZDA310）研究成果之一。
作者简介：许烨婧（ORCID：0000－0003－1128－2878），博士；黄微（ORCID：0000－0003－0448－9563），教授，博士生导师，通讯作者，E－mail：huangwei@jlu.edu.cn；赵江元（ORCID：0000－0001－5052－0598），博士研究生。
收稿日期：2020－06－04　修回日期：2020－09－09　本文起止页码：124－132　本文责任编辑：易飞

许烨婧,黄微,赵江元. 多媒体网络舆情信息并发获取模型构建及实证分析[J]. 图书情报工作,2020,64(23):124－132.

征,以上相关研究对并发处理、并行设计、系统构建、网络舆情信息获取相关的影响因素、分析方法、模型构建、信息采集等分别进行了较为详细的探讨,但这些技术和方法只是针对特定问题而展开的相关研究,有关多媒体网络舆情信息并发获取的相关研究较少,其概念、内涵尚无明确界定。相对来说,其更侧重于微观细化的针对性研究。本文将根据多媒体网络舆情相关研究热点,综合利用信息科学、统计学、管理学等基本理论,以客观现实为起点,明确并发获取的构成要素,引入多种数理分析法,通过综合分析应用,以有效确保舆情事件信息获取、分析与计算。本研究更侧重于以宏观粗粒度的视角为出发点,从复杂多变的舆情信息中获取关注度高、影响力大、波及面广的舆情信息。

本文综合上述研究并根据客观需要明确多媒体网络舆情信息并发获取的内涵,提出模型构建方法,并进行实证分析,以期在特定时间点、巨量的舆情信息中有效筛选网络舆情事件,提取关键舆情信息,为实时追踪网络舆情信息的动态变化提供强有力的客观数据支持。本研究便于管控主体及时把握舆情态势,促使相关联动主体及时调整方案,增强响应策略的科学性,避免不良舆情态势的进一步扩散和蔓延,营造持续健康和谐的舆情环境。



## 2 多媒体网络舆情信息并发获取内涵及模型构成要素

### 2.1 内涵

并发主要指单个处理器在特定时间内同时处理多个任务[13]。并行主要指多个处理器特定时间范围内同时处理多个任务[14],主要体现并行计算的能力。本研究中多媒体网络舆情信息并发获取主要是以多媒体技术为支持,并发处理海量的网络舆情信息,进行数据采集、筛选及获取的过程。

### 2.2 模型构建思路及方法

多媒体网络舆情信息并发获取模型的合理构建体现在3个方面:①解析并发获取的构成要素;②建构并发获取模型;③实证分析。主要研究方法包括管理学、统计学、模糊数学等理论:①通过文献分析法、综合调查法明晰并发获取数据源,抽象概括模型的构成要素;②结合要素分析法、层次分析法、数理分析法确定构成要素权重;③通过爬虫采集数据,并对其标准化处理,测度触发值。

### 2.3 多媒体网络舆情信息并发获取模型构成要素解析

#### 2.3.1 构成要素解析

根据网络舆情信息并发获取的数据源解析[15],本研究考虑以舆情主体、客体、媒体、本体几个维度为出发点分析其构成要素,具体如表1所示:

表1 多媒体网络舆情信息并发获取模型构成要素

| 一级构成要素 | 二级构成要素 |
| --- | --- |
| 主体参与度 | 阅读量 |
| | 讨论量 |
| | 原创量 |
| 客体热度 | 转发量 |
| | 评论量 |
| | 点赞量 |
| | 转发量均值 |
| | 评论量均值 |
| | 点赞量均值 |
| 媒体影响度 | 搜索量 |
| | 媒体量 |
| | 发布量 |
| | 日均发布量 |
| 本体融合度 | 主题词频量 |
| | 文本信息量 |
| | 主题词频均量 |
| | 文本信息均量 |

构成要素解析如下[16]:

(1)主体参与度。舆情主体参与度($SP$)是指主体阅读量(r)、讨论量(d)、原创量(o)。其中,阅读量指舆情主体某段时间内对舆情事件、话题等阅读的数量;讨论量指网络舆情主体对某舆情事件或主题讨论的次数;原创量指舆情主体采用表情、微视频等多种方式表达的原创信息数量。

(2)客体热度。文中的舆情客体热度($OH$)主要是转发量(t)、评论量(c)、点赞量(g)、转发量均值(ta)、评论量均值(ca)、点赞量均值(ga)。其中,转发量指舆情事件信息被主体转发的数量;评论量是网络舆情客体被主体点评、评论、留言的数量;点赞量是网络舆情客体被主体点赞的数量;转发量均值是单位时间内舆情事件被转发的频次;评论量均值指某舆情事件单位时间内被评论的频次;点赞量均值指舆情事件单位时间内被点赞的频次。

(3)媒体影响度。本研究中的媒体影响度($MI$)指搜索量(s)、媒体量(m)、发布量(p)、日均发布量(pa)。其中,搜索量指某段时间内检索舆情事件的频次;媒体量是某时间范围内报道某网络舆情事件的媒体数量;发布量是某时间点发布与舆情事件相关的信

息量;日均发布量指舆情事件单位时间内发布的数据流量。

（4）本体融合度。在哲学研究范畴中,本体最初被定义为“存在”的科学[17]。随着研究应用的逐步深入,其内涵更加丰富,本研究中的本体融合度（*OF*）指主题词频量（tf）、文本信息量（ti）、主题词频均量（tfa）、文本信息均量（tia）。其中,主题词频量是以舆情客体为中心所产生的舆情主题词频数量;文本信息量是舆情客体本身所体现的内容长度;主题词频均量是某时间段内围绕舆情事件所出现的本体信息中的主题词词频;文本信息均量指某段时间范围内所出现的舆情事件本体信息中的文本数量。

### 2.3.2 构成要素度量

网络舆情信息的泛在特性是各方要素共同作用的结果,其有效获取则要借助大数据处理、云计算、云存储等计算机科学技术,通过智能处理与人工识别相结合的方式采集并筛选数据,完成对各构成要素的综合度量,有益于舆情信息的快速识别与获取。通过分析现有文献信息,本研究将引用 DEMATEL、AHP、FMF 分析法进行多媒体网络舆情信息并发获取模型构建。

## 3 多媒体网络舆情信息并发获取模型构建

模型构建过程一般分为 3 个阶段:①识别网络舆情信息并发获取过程中的构成要素;②赋权予模型的各个构成要素;③测度构成要素,计算并发获取模型的触发值。如图 1 所示:

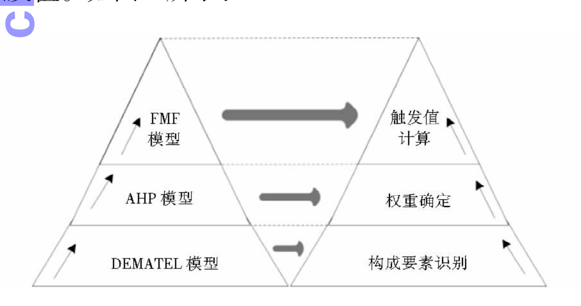



图 1 模型构建过程

### 3.1 基于 DEMATEL 的构成要素识别模型

DEMATEL 分析法是由美国 A. Gabus、E. Fontela 针对存在的复杂社会问题提出的利用专家知识、经验应对处理的方法论,能够从整体分析并识别系统内各个要素,明晰各个要素间的关系与作用。其计算步骤如下[18-20]:①构建整体系统要素,明晰主构成要素,将子构成要素记为 $E1,E2,\cdots,Em$;②研判各要素关系强弱,依据（无,弱,中,强）< = >（0,1,2,3）进行判断,明确各要素之间的影响关系;③构建系统内各要素的直接影响矩阵,标记为 $Em$;④解析要素间接影响关系,得正规化影响矩阵 $E$;⑤据公式 $T=E(I-E)-1$,计算综合影响矩阵 $T$;⑥计量各要素影响度、被影响度、中心度、原因度。

记影响度为集合 $F$,即系统内各行相对应要素对其他要素的综合影响度,$F_i$ 如公式（1）所示:

$$F_i=\sum_{j=1}^{m} t_{ij},(i=1,2,3,\cdots,m) \qquad 式（1）$$

记被影响度为集合 $D$,即系统内各列相对应要素被其他要素综合影响的作用程度,$Di$ 如公式（2）所示:

$$D_i=\sum_{j=1}^{m} t_{ji},(i=1,2,3,\cdots,m) \qquad 式（2）$$

记中心度为 $M_i$,即要素 $i$ 在整个系统的位置及其作用大小,如公式（3）所示:

$$M_i=F_i+D_i,(i=1,2,3,\cdots,m) \qquad 式（3）$$

记原因度为 $C_i$,即要素 $i$ 与其他要素之间的因果关系,如公式（4）所示:

$$C_i=F_i-D_i,(i=1,2,3,\cdots,m) \qquad 式（4）$$

最后,根据计算结果,解析系统中各要素之间的综合影响及其作用关系。

### 3.2 基于 AHP 的要素权重模型构建

AHP 主要用于多层次、多目标等复杂问题的决策,是一种将主观判断结果进行量化表示的综合决策分析法。一般要进行层次结构模型建构、判断矩阵确定、层次单排序及其一致性检验。检验标准如公式（5）（6）所示[21-22]:

$$CR=\frac{CI}{RI} \qquad 式（5）$$

$$CI=\frac{\lambda_{max}-n}{n-1} \qquad 式（6）$$

其中, RI 是平均随机一致性指标,其取值如表 2 所示:

表 2 一致性标准 RI 取值规则

| 判断矩阵阶数 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|
| *RI* | 0 | 0 | 0.58 | 0.90 | 1.12 | 1.24 | 1.32 | 1.41 | 1.45 |

当 CR < 0.1 时,则认为计算结果满足一致性检验。在此前提下,计算要素权重值。

### 3.3 基于 FMF 的网络舆情信息并发获取模型

FMF 属于模糊数学的界定范畴,在具体应用过程中,其隶属函数的确定对模糊集合具有决定作用。通过对模糊要素分析计量,模糊隶属度能够对构成要素实现系统度量,满足要素的定性分析与量化计算。FMF 已在诸多领域得到了较为广泛的研究与应用。由

此,通过模糊数学分析法计量触发值具有可操作性。步骤如下[23－24]:

(1)集结构成要素。多媒体网络舆情信息并发获取的构成要素包括4个一级主要素,17个二级子要素。触发值记为Z;一级主要素集合记为 $U=\{u_1,u_2,u_3,u_4\}$;二级子要素集合记为 $u_1=\{u_{11},u_{12},u_{13}\}$, $u_2=\{u_{21},u_{22},u_{23},u_{24},u_{25},u_{26}\}$, $u_3=\{u_{31},u_{32},u_{33},u_{34}\}$, $u_4=\{u_{41},u_{42},u_{43},u_{44}\}$。

(2)明确构成要素的隶属等级。综合测度各构成要素,建构要素的隶属等级集合。$V=\{V_1,V_2,\cdots,V_n\}$, $V_i$ 即第i个隶属等级,其标准值为 $G=\{g_1,g_2,g_3,g_4,g_5\}$,中值为 $Q=\{q_1,q_2,q_3,q_4,q_5\}$。本研究根据客观实际情况将隶属等级定为"极、强、中、一般、弱"5个等级,分别对应 $V=\{a,b,c,d,e\}$,如表3所示:

表3　构成要素量化

| 隶属等级 | 极 | 强 | 中 | 一般 | 弱 |
|---|---|---|---|---|---|
| 数值区间 | (0.8,1] | (0.6,0.8] | (0.4,0.6] | (0.2,0.4] | (0,0.2] |
| 数值(中值) | 0.9 | 0.7 | 0.5 | 0.3 | 0.1 |
| 对应表示 | a | b | c | d | e |

(3)计算各层次要素权值。

(4)建构隶属矩阵。量化二级子要素中的各要素,明确每个要素的隶属级别,建构隶属矩阵。即综合关系、舆情主体参与度、客体热度、媒体影响度、本体融合度隶属矩阵 $R$、$Ru_1$、$Ru_2$、$Ru_3$、$Ru_4$。

基于上述解析分别构建各主要素的 $r_{m*n}$ 矩阵,且:

$$r_{m*n}=\begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ r_{21} & r_{22} & \cdots & r_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ r_{m1} & r_{m2} & \cdots & r_{mn} \end{bmatrix} \quad 式(7)$$

式中, $r_{ij}(0\leqslant i\leqslant m,0\leqslant j\leqslant n)$ 表示某要素的隶属数值。

(5)计算综合测度结果。本研究引进模糊合成算子 $M(\cdot,\oplus)$,如公式(8)所示:

$$s_k=\min(1,\textstyle\sum_{j=1}^{m}\mu_j r_{jk}),(k=1,2,3,\cdots,n) \quad 式(8)$$

将构成要素 $u_1,u_2,u_3,u_4$ 二级要素各个权值 $w_i^*(i=1,2,3,4)$ 与相对应的隶属矩阵合成运算,可得各要素综合隶属值 $F_1,F_2,F_3,F_4$。

$$F_1=w_1^*\cdot(r_{u11},r_{u12},r_{u13}) \quad 式(9)$$

$$F_2=w_2^*\cdot(r_{u21},r_{u22},r_{u23},r_{u24},r_{u25},r_{u26}) \quad 式(10)$$

$$F_3=w_3^*\cdot(r_{u31},r_{u32},r_{u33},r_{u34}) \quad 式(11)$$

$$F_4=w_4^*\cdot(r_{u41},r_{u42},r_{u43},r_{u44}) \quad 式(12)$$

根据上述公式,整合计算可得单要素模糊综合测度矩阵Fu:

$$Fu=\begin{bmatrix} F_1 \\ F_2 \\ F_3 \\ F_4 \end{bmatrix} \quad 式(13)$$

依据一级要素权重 $W=(w_1,w_2,w_3,w_4)$,综合计算模糊算子与矩阵Fu,得隶属结果F,具体如公式(14)所示:

$$F=W\cdot Fu \quad 式(14)$$

(6)触发值计算。综合计量模糊测度数值与模糊隶属等级,得舆情并发获取的触发值,如公式(15)所示。

$$Z=F\cdot G^T \quad 式(15)$$

式中,Z代表隶属函数;F代表某舆情信息的模糊测度隶属值,G为舆情隶属权值。

## 4　实证分析

### 4.1　网络舆情数据源的获取

本研究以新浪微博为研究对象进行信息获取,以集搜客自动采集为主,抓取2018－2019年的网络舆情事件,共统计"俄罗斯世界杯""小米香港上市""长春长生疫苗""陕西奔驰女车主维权""重庆保时捷女车主打人""台风利奇马登录"几个事件的数据信息,将上述各事件依次简称为"Event 1"－"Event 6",其中的数据源信息涉及博主ID、博文、评论量、发布时间、转发量、讨论量、点赞量、阅读量、搜索量、发布量、原创量、文本信息量、媒体量、主题词频量等,各舆情事件因社会影响力不同,数据采集量上呈现一定区别,部分数据采集结果见图2。

在数据获取时,由于受到主客观因素的影响与限制,以上所示数据仅代表采集的部分数据,本文主要围绕抓取的数据源进行计量。

### 4.2　构成要素识别

#### 4.2.1　构成要素数据源计算

根据上述构成要素,本研究邀请若干名该研究领域专家及学者,对构成要素间的相互影响作用关系程度打分,同时根据研究需要整合打分结果,建立直接影响矩阵Em,运算正规化矩阵E、综合影响矩阵T,结果如表4所示:

| 博主 | 博主主页 | 博文 | 博文独立网址 | 发布时间 | 转发数 | 评论数 | 点赞数 | <fullpath |
|---|---|---|---|---|---|---|---|---|
| 红黑飞鱼 | https://we | #2018俄罗斯世界杯# #我爱世界杯 | https://weibo.com/1493 | 2018年06月12日 08:10 | 0 | 0 | 0 | https://s |
| 抖音娱乐局 | https://we | 倒计时2天！俄罗斯世界杯主题曲官 | https://weibo.com/2592 | 2018年06月12日 08:10 | 0 | 0 | 0 | https://s |
| 天机阁海之歌 | https://we | 海之歌说，外资开始拉美的格力了， | https://weibo.com/6451 | 2018年06月12日 08:10 | 0 | 0 | 0 | https://s |
| 环球时报 | https://we | 【倒计时3天！俄罗斯世界杯主题曲 | https://weibo.com/1974 | 2018年06月12日 08:10 | 275 | 124 | 551 | https://s |
| 江南晚报 | https://we | 【倒计时2天！俄罗斯世界杯主题曲 | https://weibo.com/1614 | 2018年06月12日 08:10 | 1 | 0 | 6 | https://s |
| 今生只恋伯纳乌 | https://we | 外国设计师根据2018俄罗斯世界杯 | https://weibo.com/1632 | 2018年06月12日 08:11 | 38 | 45 | 116 | https://s |
| 玲珑6502618748 | https://we | 俄罗斯世界杯主题起官方MV发布 好 | https://weibo.com/6502 | 2018年06月12日 08:11 | 0 | 0 | 0 | https://s |
| 新闻速播 | https://we | #世界杯来啦#【倒计时 3天！俄罗 | https://weibo.com/1873 | 2018年06月12日 08:11 | 0 | 0 | 0 | https://s |
| 寻梦人不爱做梦 | https://we | 我刚刚订阅了#世界杯#今日焦点赛 | https://weibo.com/6031 | 2018年06月12日 08:11 | 0 | 1 | 0 | https://s |
| 足球小酱w | https://we | 《转会市场》对本届世界杯最贵球 | https://weibo.com/2563 | 2018年06月12日 08:11 | 0 | 2 | 3 | https://s |
| 默声老徐 | https://we | #我是伪球迷# 哈哈哈，白岩松！！ | https://weibo.com/3259 | 2018年06月12日 08:11 | 1 | 0 | 3 | https://s |
| 少正宏朗 | https://we | 我刚刚订阅了#世界杯#今日焦点赛 | https://weibo.com/3904 | 2018年06月12日 08:11 | 0 | 1 | 0 | https://s |
| 陌上有天人如玉 | https://we | #欢乐竞猜# 2018年俄罗斯世界杯 | https://weibo.com/6493 | 2018年06月12日 08:11 | 0 | 0 | 0 | https://s |
| T丶Spring | https://we | 我刚刚订阅了#世界杯#今日焦点赛 | https://weibo.com/5385 | 2018年06月12日 08:12 | 0 | 1 | 0 | https://s |
| 陕西都市快报 | https://we | 【收藏丨2018世界杯赛程表 拿走不 | https://weibo.com/1915 | 2018年06月12日 08:12 | 15 | 7 | 12 | https://s |
| 安徽工程大学人 | https://we | #大千世界#【倒计时 3天！俄罗斯 | https://weibo.com/2683 | 2018年06月12日 08:12 | 0 | 0 | 0 | https://s |
| 江淮气象 | https://we | #喝彩世界杯# 俄罗斯世界杯就要开 | https://weibo.com/2240 | 2018年06月12日 08:12 | 2 | 6 | 10 | https://s |
| 上海黄浦 | https://we | 倒计时3天！俄罗斯世界杯主题曲官 | https://weibo.com/2557 | 2018年06月12日 08:12 | 0 | 0 | 1 | https://s |
| 其美南加turbo | https://we | 我刚刚订阅了#世界杯#今日焦点赛 | https://weibo.com/5354 | 2018年06月12日 08:12 | 0 | 0 | 0 | https://s |
| 用户6573057871 | https://we | 2018年俄罗斯世界杯主题曲（ 来自 | https://weibo.com/6573 | 2018年06月12日 08:12 | 0 | 0 | 0 | https://s |
| 足球小酱w | https://we | 迪巴拉的射门了解下，踢个死角而 | https://weibo.com/2563 | 2018年06月12日 08:12 | 0 | 1 | 2 | https://s |
| 牛镜看互联网 | https://we | 【早安，一起放飞自我】倒计时三 | https://weibo.com/6371 | 2018年06月12日 08:13 | 0 | 1 | 0 | https://s |
| 南京广播电视台 | https://we | 【倒计时3天！俄罗斯世界杯主题曲 | https://weibo.com/2663 | 2018年06月12日 08:13 | 2 | 0 | 2 | https://s |

图 2　俄罗斯世界杯数据源截图

表 4　综合影响矩阵 T

| 序号 | E1 | E2 | E3 | E4 | E5 | E6 | E7 | E8 | E9 | E10 | E11 | E12 | E13 | E14 | E15 | E16 | E17 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| E1 | 0.022 | 0.048 | 0.039 | 0.072 | 0.092 | 0.093 | 0.042 | 0.043 | 0.043 | 0.082 | 0.013 | 0.023 | 0.017 | 0.017 | 0.019 | 0.013 | 0.014 |
| E2 | 0.052 | 0.036 | 0.049 | 0.083 | 0.082 | 0.083 | 0.052 | 0.051 | 0.051 | 0.050 | 0.020 | 0.054 | 0.048 | 0.049 | 0.051 | 0.042 | 0.044 |
| E3 | 0.046 | 0.051 | 0.021 | 0.076 | 0.075 | 0.075 | 0.045 | 0.045 | 0.045 | 0.065 | 0.015 | 0.047 | 0.021 | 0.042 | 0.044 | 0.017 | 0.018 |
| E4 | 0.073 | 0.100 | 0.067 | 0.068 | 0.106 | 0.105 | 0.111 | 0.069 | 0.069 | 0.066 | 0.053 | 0.115 | 0.105 | 0.088 | 0.090 | 0.057 | 0.060 |
| E5 | 0.064 | 0.092 | 0.060 | 0.099 | 0.057 | 0.097 | 0.063 | 0.102 | 0.062 | 0.060 | 0.047 | 0.086 | 0.078 | 0.060 | 0.082 | 0.051 | 0.054 |
| E6 | 0.066 | 0.070 | 0.042 | 0.075 | 0.075 | 0.035 | 0.045 | 0.044 | 0.085 | 0.043 | 0.034 | 0.045 | 0.021 | 0.020 | 0.022 | 0.015 | 0.017 |
| E7 | 0.056 | 0.062 | 0.053 | 0.109 | 0.067 | 0.066 | 0.035 | 0.054 | 0.054 | 0.053 | 0.043 | 0.079 | 0.052 | 0.054 | 0.055 | 0.045 | 0.047 |
| E8 | 0.055 | 0.082 | 0.051 | 0.087 | 0.107 | 0.087 | 0.054 | 0.033 | 0.053 | 0.052 | 0.021 | 0.057 | 0.051 | 0.051 | 0.053 | 0.044 | 0.046 |
| E9 | 0.044 | 0.047 | 0.041 | 0.052 | 0.051 | 0.093 | 0.043 | 0.043 | 0.023 | 0.021 | 0.014 | 0.044 | 0.039 | 0.020 | 0.041 | 0.015 | 0.037 |
| E10 | 0.075 | 0.080 | 0.050 | 0.086 | 0.085 | 0.085 | 0.053 | 0.052 | 0.052 | 0.031 | 0.041 | 0.055 | 0.028 | 0.049 | 0.051 | 0.042 | 0.044 |
| E11 | 0.108 | 0.116 | 0.083 | 0.125 | 0.124 | 0.082 | 0.087 | 0.065 | 0.063 | 0.065 | 0.027 | 0.070 | 0.061 | 0.083 | 0.086 | 0.052 | 0.055 |
| E12 | 0.141 | 0.153 | 0.133 | 0.168 | 0.165 | 0.164 | 0.137 | 0.135 | 0.135 | 0.114 | 0.090 | 0.082 | 0.127 | 0.132 | 0.136 | 0.114 | 0.120 |
| E13 | 0.135 | 0.146 | 0.127 | 0.160 | 0.157 | 0.156 | 0.132 | 0.130 | 0.130 | 0.088 | 0.086 | 0.137 | 0.061 | 0.126 | 0.130 | 0.090 | 0.095 |
| E14 | 0.127 | 0.137 | 0.100 | 0.148 | 0.127 | 0.125 | 0.104 | 0.101 | 0.101 | 0.102 | 0.081 | 0.128 | 0.096 | 0.059 | 0.123 | 0.085 | 0.090 |
| E15 | 0.076 | 0.084 | 0.069 | 0.112 | 0.110 | 0.109 | 0.073 | 0.071 | 0.071 | 0.110 | 0.056 | 0.098 | 0.086 | 0.111 | 0.052 | 0.079 | 0.102 |
| E16 | 0.117 | 0.148 | 0.128 | 0.160 | 0.158 | 0.156 | 0.113 | 0.110 | 0.110 | 0.128 | 0.107 | 0.137 | 0.122 | 0.128 | 0.112 | 0.049 | 0.095 |
| E17 | 0.106 | 0.115 | 0.098 | 0.126 | 0.123 | 0.122 | 0.102 | 0.100 | 0.100 | 0.100 | 0.081 | 0.127 | 0.094 | 0.119 | 0.122 | 0.085 | 0.048 |



根据公式(1)－(4)分别计算构成要素的影响度、被影响度、中心度、原因度，结果见表 5。

### 4.2.2　构成要素数据源计算结果分析

根据表 5 可知，各构成要素关联关系密切。通常中心度、原因度相对而言更受认可。因此，本研究的构成要素识别及其要素间的相互作用影响程度主要从这两个视角展开探讨。

(1)中心度分析。中心度值越大，要素所起作用就越大，其影响程度也越大。根据表 5，E4、E12、E13、E14 与其他要素相比，其中心度测度值更大，综合影响度更强；E1、E2、E5、E6、E7、E8、E10、E11、E15、E16、E17 的中心度测度值大小居中，表明仍对其他要素构成重大影响；E3、E9 的中心度测度值与其他要素相比略偏小，其测度值结果仍在整个体系内，没有呈现大幅偏离现象。依据中心度结果，各构成要素密不可分，相互作用强度较大。

(2)原因度分析。通过计算可知 E11、E12、E13、E14、E15、E16、E17 的测度结果为正，是系统中的原因要素，其影响度更大。相对的，E1、E2、E3、E4、E5、E6、E7、E8、E9、E10 的测度结果为负，则是系统的结果要素，其影响度相对较弱。据此可明晰系统各要素均发挥着一定程度的影响与作用。

综观上述各构成要素间的影响程度，从总体上看其数据测度计算结果较为聚集，虽存在局部影响关系

表 5　构成要素综合影响指数

| 序号 | 影响度 | 被影响度 | 中心度 | 原因度 |
|---|---|---|---|---|
| E1 | 0.693 | 1.362 | 2.055 | -0.670 |
| E2 | 0.897 | 1.569 | 2.466 | -0.672 |
| E3 | 0.748 | 1.211 | 1.959 | -0.463 |
| E4 | 1.403 | 1.807 | 3.210 | -0.404 |
| E5 | 1.212 | 1.761 | 2.973 | -0.550 |
| E6 | 0.754 | 1.732 | 2.486 | -0.978 |
| E7 | 0.983 | 1.290 | 2.273 | -0.308 |
| E8 | 0.984 | 1.249 | 2.234 | -0.265 |
| E9 | 0.669 | 1.248 | 1.917 | -0.580 |
| E10 | 0.958 | 1.228 | 2.186 | -0.271 |
| E11 | 1.353 | 0.829 | 2.182 | 0.524 |
| E12 | 2.247 | 1.385 | 3.632 | 0.862 |
| E13 | 2.087 | 1.106 | 3.194 | 0.981 |
| E14 | 1.833 | 1.206 | 3.039 | 0.626 |
| E15 | 1.469 | 1.270 | 2.739 | 0.199 |
| E16 | 2.078 | 0.892 | 2.970 | 1.186 |
| E17 | 1.767 | 0.986 | 2.753 | 0.781 |

略弱的现象,但仍在整个系统内存在影响,且对其他构成要素存在相互作用关系。由此,本研究将上述要素综合纳入探究范畴,共同构成了多媒体网络舆情信息并发获取不可或缺的重要要素。

### 4.3　构成要素权重确定

依据公式(5)(6)计算可得一级要素 $CR_W$ = 0.073 358;二级要素 $CRw_p$ = 0.033 199, $CRw_h$ = 0.089 587, $CRw_i$ = 0.043 327, $CRw_f$ = 0.073 358。CR 值小于 0.1,通过一致性检验。

计算得一、二级构成要素的权重分别为:

$$W = \begin{bmatrix} 0.151\ 18 \\ 0.508\ 29 \\ 0.075\ 2 \\ 0.265\ 34 \end{bmatrix}, \quad Wp = \begin{bmatrix} 0.636\ 99 \\ 0.258\ 28 \\ 0.104\ 73 \end{bmatrix},$$

$$Wh = \begin{bmatrix} 0.415\ 46 \\ 0.229\ 44 \\ 0.143\ 17 \\ 0.113\ 3 \\ 0.063\ 232 \\ 0.035\ 406 \end{bmatrix}, \quad Wi = \begin{bmatrix} 0.117\ 5 \\ 0.055\ 285 \\ 0.565\ 01 \\ 0.262\ 2 \end{bmatrix},$$

$$Wf = \begin{bmatrix} 0.508\ 29 \\ 0.265\ 34 \\ 0.151\ 18 \\ 0.075\ 2 \end{bmatrix}$$

上述各层次构成要素权重值的确定对计算触发值具有关键作用。

### 4.4　触发值计算

(1)构成要素集结。一级构成要素 U = {SP, OH, MI, OF};二级构成要素 $u_1$ = {r,d,o}, $u_2$ = {t,c,g,ta,ca,ga}, $u_3$ = {s,m,p,pa}, $u_4$ = {tf,ti,tfa,tia}。

(2)明确构成要素隶属等级及数据标准化。根据表 3,V = {弱,一般,中,强,极},标准值 G = (0.2,0.4,0.6,0.8,1),根据数据源获取状况,对数据标准化处理[25],为构建隶属关系矩阵提供规范化数值。

(3)量化隶属矩阵。量化各要素隶属等级,建构标准化隶属矩阵[26]。方法如下:

$$v_a(n_{ij}) = \begin{cases} 0 & n_{ij} < b \\ \dfrac{n_{ij}-b}{a-b} & n_{ij} \in [b,a) \\ 1 & n_{ij} \geqslant a \end{cases} \qquad 式(16)$$

$$v_b(n_{ij}) = \begin{cases} 0 & n_{ij} < c, \text{or } n_ij \geqslant a \\ \dfrac{n_{ij}-c}{b-c} & n_{ij} \in [c,b) \\ \dfrac{a-n_{ij}}{a-b} & n_{ij} \in [b,a) \end{cases} \qquad 式(17)$$

$$v_c(n_{ij}) = \begin{cases} 0 & n_{ij} < d, \text{or } n_{ij} \geqslant b \\ \dfrac{n_{ij}-d}{c-d} & n_{ij} \in [d,c) \\ \dfrac{b-n_{ij}}{b-c} & n_{ij} \in [c,b) \end{cases} \qquad 式(18)$$

$$v_d(n_{ij}) = \begin{cases} 0 & n_{ij} < e, \text{or } n_{ij} \geqslant c \\ \dfrac{n_{ij}-e}{d-e} & n_{ij} \in [e,d) \\ \dfrac{c-n_{ij}}{c-d} & n_{ij} \in [d,c) \end{cases} \qquad 式(19)$$

$$v_e(n_{ij}) = \begin{cases} 0 & n_{ij} \geqslant d \\ \dfrac{d-n_{ij}}{d-e} & n_{ij} \in [e,d) \\ 1 & n_{ij} < e \end{cases} \qquad 式(20)$$

其中,$n_{ij}$为规范化的构成要素数值,参照公式(16)-(20),定量计算各层次要素隶属级别。据公式(7)分别列出构成要素的舆情主体参与度、客体热度、媒体影响度、本体融合度隶属矩阵。

(4)构成要素综合测度。参照公式(8)-(12),对标准化的数据进行测度,可知各事件的单要素模糊隶属值。

参照公式(14),综合运算 W 与 Fu,可得各事件综合测度隶属结果 F,具体如下:

Event 1 F: F = W · Fu = [0.724 067　0.108 156

0.011 859　0.060 748 399　0.095 186]

Event 2 F:F = W · Fu = [0　0.020 737　0.089 370 7　0.322 750 2　0.567 157 9]

Event 3 F:F = W · Fu = [0.214 622 2　0.197 706 8　0.273 888 5　0.309 782 4　0.004 016 5]

Event 4 F:F = W · Fu = [0.107 875 6　0.107 459 5　0.067 110 6　0.259 707 1　0.457 863 5]

Event 5 F:F = W · Fu = [0.201 908 2　0.038 657 3　0　0.116 351 3　0.643 099 5]

Event 6 F:F = W · Fu = [0.205 713 9　0.148 664 7　0.208 832 9　0.081 249 5　0.428 327 2]

(5)触发值计算。依据公式(15)计算触发值,各事件测度结果如下:

Event 1:$Z = F \cdot G^T = 0.861$;Event 2:$Z = F \cdot G^T = 0.313$;Event 3:$Z = F \cdot G^T = 0.662$;Event 4:$Z = F \cdot G^T = 0.430$;Event 5:$Z = F \cdot G^T = 0.408$;Event 6:$Z = F \cdot G^T = 0.568$。

综上,“Event 1” - “Event 6”的并发获取触发值结果如图 3 所示:

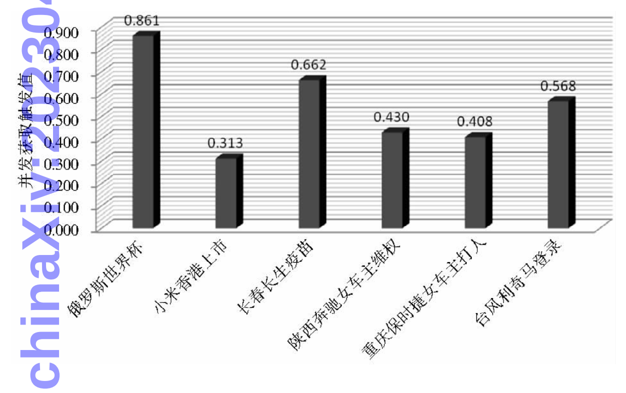



图 3　网络舆情事件并发获取触发值

根据图 3 可知各网络舆情事件的并发获取触发值,依照触发值结果大小及其标号分别将其排序,结果为“Event 1”“Event 3”“Event 6”“Event 4”“Event 5”“Event 2”。

## 4.5　实验结果解析

根据上述构建的多媒体网络舆情信息并发获取模型,通过对各网络舆情事件数据的获取与处理,参照测度结果,本研究将从事件的隶属等级层次划分与并发获取触发值视阈展开深度解析。

### 4.5.1　事件隶属等级层次划分

本文根据现实研究需要将舆情事件的隶属等级划分为 4 个层次:①若网络舆情事件的隶属等级为极、强,则说明事件本身已经大面积覆盖众多区域,网络舆情主体在不同的时空、区域体现出极强的活跃参与度,在特定的时空环境造成极大影响,且伴随时间变化飞速扩散。此时,直接唤起触发状态,开启衍进追踪。若舆情事件反响积极,则宣扬、鼓励;反之,则实施监测追踪,根据数据追踪态势,及时制定较为合理的响应决策,积极引导、释疑、疏解,在此期间实时洞悉网络舆情事件的衍进变化,总结应对工作中存在的问题与不足。②若网络舆情事件的隶属等级为中,则从两个方面考虑是否进一步展开舆情话题衍进追踪。其一,若触发值高于区间中值,则启动衍进追踪;其二,若低于区间中值,则停止获取舆情信息,同时也可参照网络舆情事件的具体属性,进而决定是否展开进一步的信息获取与衍进追踪。③若网络舆情事件的隶属等级一般,通常情况不需开启触发模式,但相关管控主体可根据舆情事件的本质、特性等客观情况以及实际需要做出是否开启触发模式的决定。④若网络舆情事件的隶属等级为弱,一般可以忽略不计,无需开启触发模式,适时监测即可。网络舆情事件“Event 1” - “Event 6”隶属等级划分见表 6。

从表 6 中可以看出,Event 1 的触发值最大,级别最高,位于极值隶属区间;Event 3 的计量数值结果次之,隶属等级为强,其相对来说较接近极值区间;Event 6 的隶属等级为中,其较接近强值隶属区间;Event 4 的隶属等级为中;Event 5 的隶属等级为中;Event 2 的隶属等级为一般。

### 4.5.2　触发结果分析

根据上述分析可见,隶属等级属于极、强区间的有 Event 1、Event 3。隶属等级为中的事件有 Event 6、Event 4、Event 5。隶属等级一般的事件为 Event 2。其中,Event 1 是热点国际体育赛事,宣扬运动、拼搏等精神,有鉴于此,可给予宣传、支持,开启触发模式,展开实时动态追踪,分析数据走势的衍进特性,传播扩散正能量。Event 3 是医疗卫生领域热度较高的话题之一,该事件自出现时,即引发了广泛关注,产生了消极不良影响,管控主体十分重视。据此应启动触发模式,多视角开展话题衍进追踪,根据追踪态势实时制定合情合理的响应策略,问责、调查、处置并追究涉事主体的法律责任,时刻关注网民的观点、态度、情绪,并及时引导、回应、疏解,提出合理的解决方案,避免民众负面情绪的加剧以及事态的进一步蔓延。Event 6 涉及自然灾害话题范畴,具有突发性、随机性,其结果值处于中级区间,且高于区间中值,达到触发状态,可启动衍进追踪。管控主体要实时监测、追踪, 根据统计测算结

表6　网络舆情事件并发获取数值结果

| 网络舆情事件 | 俄罗斯世界杯（Event 1） | 小米香港上市（Event 2） | 长春长生疫苗（Event 3） | 陕西奔驰女车主维权（Event 4） | 重庆保时捷女车主打人（Event 5） | 台风利奇马登录（Event 6） |
|---|---|---|---|---|---|---|
| 触发值 | 0.861 | 0.313 | 0.662 | 0.430 | 0.408 | 0.568 |
| 隶属级别 | 极 | 一般 | 强 | 中 | 中 | 中 |
| 是否获取 | 是 | 否 | 是 | 否 | 否 | 是 |

果,有针对地采取各项措施,增强大众防范意识并提高其安全感,保持良性态势。Event 4、Event 5 属于社会性话题,其等级均位于中级区间,但其触发值低于中值,无需衍进追踪。主题上,两个事件存在共性因素,与车有关,话题邻近,其触发值也较接近,当然,管控主体可根据实情做出相应的决定。Event 2 与电子科技话题有关,其结果值相对偏低,是民生的体现,管控主体可据情响应。

通过分析可见,Event 1、Event 3、Event 6 都符合触发状态,均可对其实行衍进追踪。Event 4、Event 5 则处于中级隶属区间,Event 2 属于一般级别,在本研究中,均不需启动触发状态,管控主体可根据客观情况灵活选择是否衍进追踪。

经过对比发现,本实验案例分析与现实网络舆情信息演进变化较为接近,一定程度上能够客观反映网络舆情信息获取状况,基于网络舆情信息并发获取模型的构建,一方面有利于重大网络舆情事件的快速有效获取,有助于管控主体研判网络舆情信息的走势变化,及时做好响应策略,预防并降低潜在的网络舆情危机;另一方面有利于对该网络舆情事件衍进追踪,实时监测舆情信息,及时调整对应方案,营造健康网络舆情环境。该环节能够有效发现具有影响力的网络舆情事件,为网络舆情实时分析提供客观的数据支持。

## 5　结语

在大量文献调查与分析的基础上,本文综述了当前网络舆情信息获取的研究现状,界定了多媒体网络舆情信息并发获取的内涵并选取了其构成要素,明确了模型构建过程,基于 DEMATEL、AHP、FMF 分析法构建并发获取模型,采集事件数据进行实证分析。研究结果表明,本研究提出的理念与方法对重要网络舆情事件的筛选作用明显,提供的信息较为客观,在研究思路、模型构建、数据分析上均具有创新性,为多媒体网络舆情信息并发获取的理论与实践研究提供了有效的参考。但也存在一定的局限:本文的模型构建与现有的模型数据对比而言,由于研究对象等诸多主客观因素影响,存在一定的差异性,尚不能进行精确对比分析,以上研究仅是尝试性探索分析,尚有很大的挖掘空间,框架体系还有待补充完善,未来可针对不同主体、行业领域、机构等需求展开个性化的信息获取服务。下一步的研究中将增加数据的对比分析,可尝试采用其他邻近学科的研究理念与方法展开多视角的探究。

**参考文献:**

[1] 中国互联网络信息中心. 中国互联网络发展状况统计报告[EB/OL]. [2020－05－10]. http://www.cnnic.net.cn/hlwfzyj/hlwxzbg/hlwtjbg/202004/t20200428_70974.htm.
[2] XUE J, YANG Z, HOU S, et al. Processing concurrent graph analytics with decoupled computation model[J]. IEEE transactions on computers, 2017, 66(5):876－890.
[3] ZHENG H, LIU W, XIAO C. Structural relationship model for design defect and influencing factors in the concurrent design process [J]. International journal of production research, 2018, 56(13/14):4897－4924.
[4] SUN S, GONG J, ZOMAYA A Y, et al. A distributed incremental information acquisition model for large-scale text data[J]. Cluster computing, 2019,22(1):2383－2394.
[5] LIU Y, KANG X, MIAO D, et al. A knowledge acquisition method based on concept lattice and inclusion degree for ordered information systems[J]. International journal of machine learning and cybernetics, 2019,10(1):3245－3261.
[6] LIU J. Information collection principles, main channels and common methods of public sector websites[M]// HUNG J C, YEN N Y, CHANG J W. Frontier computing. Singapore: Springer,2020:518－524.
[7] 李辉,王建文,叶明雯. 基于 Hadoop 的海量气象水文数据并发处理模型[J]. 计算机应用,2018,38(S2):187－191,205.
[8] 洪亮,石立艳,李明. 基于系统动力学的多主体回应网络舆情影响因素研究[J]. 情报科学,2017,35(1):133－138.
[9] 朱红灿, 胡新, 廖小巧. 基于心流理论的公众政府信息获取网络渠道持续使用意愿研究[J]. 情报资料工作, 2018,39(2):56－62.
[10] 靳江艳, 方忆湘, 周京博, 等. 坐标测量信息多色集合层次结构表示及获取[J]. 组合机床与自动化加工技术, 2019, 540(2):65－67,71.
[11] 李莉,刘淼,冯嘉辉. 基于多元数据信息获取的关键技术研究[J]. 长春理工大学学报(自然科学版),2020,43(1):97－103.
[12] 张洪豪,赵泽宁,姚世春. 一种基于 Web 爬虫的网页信息获取系统[J]. 中国科技信息,2020,624(6):90－91.

[13] 秦宜州. 实时监控系统中基于多核并发的并行技术应用研究[D]. 沈阳:东北大学,2012.

[14] GAO Q, QIN F, PANDA D K. DMTracker: finding bugs in large-scale parallel programs by detecting anormaly in data movements[C]//IEEE conference on supercomputing. Reno, Nevada: Association for Computing Machinery,2007.

[15] 黄微,许烨婧,刘熠. 大数据环境下多媒体网络舆情并发获取的数据驱动机理研究[J]. 情报理论与实践,2019,42(6):42 - 48,16.

[16] 许烨婧. 多媒体网络舆情信息的并发获取机理与话题衍进追踪研究[D]. 长春:吉林大学,2020.

[17] WC recommendation. OWL Web ontology language overview[J]. February, 2004, 63(45):990 - 996.

[18] 杨印生,李洪伟. 管理科学与系统工程中的定量分析方法[M]. 长春:吉林科学技术出版社,2009.

[19] 李宗富,张向先. 政务微信公众号服务质量的关键影响因素识别与分析[J]. 图书情报工作,2016,60(14):84 - 93.

[20] 许鹏程,毕强,丁梦晓,等. 基于 DEMATEL 的知识发现系统交互质量影响因素识别与分析[J]. 情报资料工作,2018,39(4):24 - 31.

[21] GÜNDODU F K, KAHRAMAN C. A novel spherical fuzzy analytic hierarchy process and its renewable energy application[J]. Soft computing, 2020, 24(6):4607 - 4621.

[22] 王高飞,李明. 基于 AHP - 模糊综合分析的移动社交网络舆情预警模型研究[J]. 现代情报,2017,37(1):41 - 44,65.

[23] 相甍甍. 移动电子商务消费者隐私信息披露行为及风险研究[D]. 长春:吉林大学,2018.

[24] 黄微,许烨婧,韩瑞雪,等. 网络舆情语义倾向性的隶属度研究[J]. 图书情报工作,2015,59(21):27 - 32.

[25] 陈金强. 指标权重信息不完全下的模糊综合评价方法及其在绩效考核中的应用研究[D]. 杭州:浙江财经大学,2017.

[26] 张国力. 基于模糊综合评价的国有大型企业信息安全管理研究[D]. 北京:北京邮电大学,2012.

作者贡献说明:
许烨婧:论文整体撰写与修改;
黄微:研究命题及研究思路的制订,论文指导及最后版本修订;
赵江元:论文数据搜集与整理。



## The Model Construction and Empirical Analysis on Concurrent Acquisition of Multimedia Network Public Opinion Information

Xu Yejing　Huang Wei　Zhao Jiangyuan
School of Management, Jilin University, Changchun 130022

Abstract: [**Purpose/significance**] The model construction and empirical study on concurrent acquisition of network public opinion information are helpful to obtain the key information needed in time from the massive public opinion information and provide data guarantee for its effective analysis. [**Method/process**] Through the comprehensive analysis on the current situation of public opinion information acquisition research, this paper defined the model elements of concurrent acquisition of multimedia network public opinion information, and constructed the model by combining three mathematical analysis methods of DEMATEL, AHP and FMF, and carried out empirical analysis accordingly. [**Result/conclusion**] The research results show that the data conclusions obtained are more consistent with the objective situation of public opinion events, which can be used as the basis for the judgment of concurrent acquisition of public opinion information.

Keywords: multi-media　network public opinion　concurrent acquisition　model construction